Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON THE ERRORS INCIDENT TO VALUES INTERPOLATED BY MEANS OF FIRST DIFFERENCES FROM TABLES OF LOGARITHMS, NATURAL TRIGONOMETRIC FUNCTIONS, ETC.

By R. S. Woodward, Washington, D. C.

The theory of the errors incident to values interpolated by means of first differences from a table of logarithms, natural sines, etc., appears to be a matter of considerable difficulty. Our text-books are notably deficient in information on this subject or notably erroneous in such information as they give; indeed few writings on the subject are free from error.

The most common departure from correct procedure in treating this class of errors is in assuming the applicability to them of the law of error of least squares. In addition, the magnitude and continuity or discontinuity of the errors involved are frequently misunderstood. Here also as in other branches of the theory of errors, the actual, probable, mean, and average errors are apt to be confounded with one another.

Since attention has recently been called in the Annals by Prof. Howe to the errors incident to the most frequently employed species of interpolated values, it may be of interest to point out some of the essential features of the correct theory of such errors.

In a paper by the writer published in *The Analyst*, 1882, four species of interpolated values were considered. Two of these depended on the tabular differences and two on the true differences. It was shown that the latter are less precise than the former, or, paradoxical as it may seem, that the use of the true differences is less advantageous than the use of the tabular differences. Of the other two species, the first retains one or more figures beyond the tabular number of figures, and the second (which is most commonly employed) rounds the interpolated value to the nearest unit of the last tabular place. It will be convenient to refer to these two as the the first and second species.

The theory of the errors of the first species has been fully worked out by Bessel (*Astronomische Nachrichten*, Nos. 358 and 359, 1838) and affords a basis for the theory of the errors of the second species.

I am not aware that the theory of the errors of the second class has been completely developed. In the paper referred to above, however, I have indicated the process of deriving the laws of frequency of the errors of the second species and have given the laws and consequent probable errors for several special cases. I am now able to extend the process and give more general results.

The actual error ε_0 of an interpolated value of the first species is expressed by the equation

$$\varepsilon_0 = (1 - t)\,\varepsilon_1 + t\varepsilon_2, \qquad (1)$$

in which ε_1 and ε_2 are the actual errors of the tabular logarithms or numbers, and t is the interpolating factor, which varies from 0 to 1. ε_1 and ε_2 are expressed in units of the last tabular place and may have any value from $-\frac{1}{2}$ to $+\frac{1}{2}$; hence also ε_0 may have any value between the limits $-\frac{1}{2}$ and $+\frac{1}{2}$.

The actual error ε of an interpolated value of the second species will equal that for the first species plus the error which comes from rounding the product of the interpolating factor and the tabular difference to the nearest unit. Hence, if we call the error from the latter source ε_3, we shall have

$$\varepsilon = (1 - t)\,\varepsilon_1 + t\varepsilon_2 + \varepsilon_3. \qquad (2)$$

It must now be observed that ε_3 differs from ε_1 and ε_2 in two important respects, viz: 1°, it is discontinuous, and 2°, it is dependent on t. Thus for $t = \frac{1}{2}$, ε_3 must be either 0 or $\frac{1}{2}$. Again, for $t = \frac{1}{3}$, ε_3 is restricted to one of three values, viz: 0, $+\frac{1}{3}$ and $-\frac{1}{3}$. On this account the maximum value of ε will vary also with t. For $t = \frac{1}{2}$ the maximum value of $\varepsilon = \frac{1}{2} + \frac{1}{2} = 1$, but for $t = \frac{1}{3}$ the maximum of $\varepsilon = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$; this maximum cannot always be 1 as some of the books assert. Moreover, if t is small relatively to the largest tabular difference for the table used, ε_3 will be correspondingly small. Thus, in a 5-place table of logarithms, the largest tabular difference is 44; and if $t = 0.001$ the maximum value of ε_3 is 0.044 and the maximum of ε is 0.544.*

To determine the probable, mean, and average errors corresponding to ε_0 and ε_1 we must know the laws of frequency of these actual errors. If we represent the probability of the occurrence of an error ε by $\varphi(\varepsilon)\,d\varepsilon$, it may be shown that† for the system of errors represented by equation (1)

$$\begin{aligned}\varphi(\varepsilon_0) &= \frac{\frac{1}{2} + \varepsilon_0}{(1 - t)t} \text{ for values of } \varepsilon_0 \text{ between } -\tfrac{1}{2} \text{ and } -(\tfrac{1}{2} - t),\\ &= \frac{1}{1 - t} \text{ for values of } \varepsilon_0 \text{ between } -(\tfrac{1}{2} - t) \text{ and } +(\tfrac{1}{2} - t), \qquad (3)\\ &= \frac{\frac{1}{2} - \varepsilon_0}{(1 - t)\,t} \text{ for values of } \varepsilon_0 \text{ between } +(\tfrac{1}{2} - t) \text{ and } +\tfrac{1}{2}.\end{aligned}$$

As our present purpose is to discuss the errors of the second species, we shall not derive the probable and other critical errors of the first species. It is sufficient to remark that these critical errors show that an interpolated value of the first species is in general much more precise than a tabular value.†

To derive the laws of frequency of the errors ε of the second species, it is necessary to combine the continuous errors ε_0, whose laws of frequency are

---

[*] These facts would seem to be perfectly plain, but experience justifies the assertion that they are not commonly appreciated.

[†] See Bessel's paper or that of the writer referred to above.

given by (3), with the discontinuous errors ε_3. First consider the case in which t is a fraction of the form

$$t = \frac{1}{n},$$

in which n is an integer.* Then the possible values of ε_3 are the following : —

$$\left.\begin{array}{l} 0, +\frac{1}{n}, +\frac{2}{n}, \ldots\ldots +\frac{n-1}{2n} \\ -\frac{1}{n}, -\frac{2}{n}, \ldots\ldots -\frac{n-1}{2n} \end{array}\right\} \text{for } n \text{ odd,}$$

$$\left.\begin{array}{l} 0, +\frac{1}{n}, +\frac{2}{n}, \ldots\ldots +\frac{n-2}{2n} \\ -\frac{1}{n}, -\frac{2}{n}, \ldots\ldots -\frac{n-2}{2n} \\ \qquad\qquad\qquad\quad \pm\frac{1}{2} \end{array}\right\} \text{for } n \text{ even.}$$

These errors are all equally likely to occur. The sign of the error $\frac{1}{2}$ for n even is arbitrary. It will here be considered positive only; and, in deriving the laws of frequency of the errors ε for n even, they will be considered with respect to magnitude simply, or as if they all had the same sign. It will be observed that the numerical maximum of ε is

$$\pm\left(\frac{1}{2} + \frac{n-1}{2n}\right) = \pm\frac{2n-1}{2n} \text{ for } n \text{ odd, and}$$

$$+\tfrac{1}{2} + \tfrac{1}{2} = +1 \text{ for } n \text{ even.}$$

Now if we add each of the above n values of ε_3 to ε_0 in (3), we shall have n sets of probability functions, $\varphi(\varepsilon_0 + \varepsilon_3) = \varphi(\varepsilon)$. If then the element-areas $\varphi(\varepsilon)\,d\varepsilon$ defined by these several functions be summed for every value of ε, the result will be the probability area for the whole system of errors. The area enclosed by the straight lines whose equations are given by (3) and the axis of ε_0, is 1. Hence the sum of n such unit areas is n, and we must divide the results of the above summation by n in order to have the integral of $\varphi(\varepsilon)\,d\varepsilon$ over the whole range of error, or the sum of the probabilities of all possible errors ε, unity. Without going further into the details of this process, which may be easily illustrated and worked out graphically, the results are as follows : —

For n odd

$$\varphi(\varepsilon) = \frac{n}{n-1}\left(\frac{2n-1}{2n} + \varepsilon\right) \text{ for values of } \varepsilon \text{ between } -\frac{2n-1}{2n} \text{ and } -\frac{1}{2n}, \quad (4)$$

---

*For a given table n must not exceed twice the greatest tabular difference.

$$= 1 \qquad \text{for values of } \varepsilon \text{ between } -\frac{1}{2n} \text{ and } +\frac{1}{2n}, \qquad (4\ \textit{cont.})$$

$$= \frac{n}{n-1}\left\{\frac{2n-1}{2n} - \varepsilon\right\} \quad \text{for values of } \varepsilon \text{ between } +\frac{1}{2n} \text{ and } +\frac{2n-1}{2n}.$$

For n even

$$\varphi(\varepsilon) = \frac{n}{n-1}\left\{\frac{2(n-1)}{n} - \varepsilon\right\} \quad \text{for values of } \varepsilon \text{ between } 0 \text{ and } \frac{1}{n},$$

$$= \frac{2n}{n-1}\left\{\frac{2n-1}{2n} - \varepsilon\right\} \quad \text{for values of } \varepsilon \text{ between } \frac{1}{n} \text{ and } \frac{n-1}{n}, \qquad (5)$$

$$= \frac{n}{n-1}(1-\varepsilon) \qquad \text{for values of } \varepsilon \text{ between } \frac{n-1}{n} \text{ and } 1.$$

For n odd the probability area is a trapezoid symmetrical with respect to the axis of $\varphi(\varepsilon)$. For n even the area is bounded by three straight lines and the positive axes of ε and $\varphi(\varepsilon)$, the errors as stated above being considered with respect to magnitude only.

Now let ε_p = the probable error,
ε_m = the mean error,
ε_a = the average error.

The probable error is derived from the condition

$$\int_0^{\pm\varepsilon_p} \varphi(\varepsilon)\, d\varepsilon = \tfrac{1}{4} \text{ for } n \text{ odd,}$$

$$= \tfrac{1}{2} \text{ for } n \text{ even.}$$

The square of the mean error

$$\varepsilon_m{}^2 = 2\int_0^{\frac{2n-1}{2n}} \varphi(\varepsilon)\,\varepsilon^2\, d\varepsilon \quad \text{for } n \text{ odd,}$$

$$= \int_0^{1} \varphi(\varepsilon)\,\varepsilon^2 d\varepsilon \qquad \text{for } n \text{ even.}$$

The average error

$$E_a = 2\int_0^{\frac{2n-1}{2n}} \varphi(\varepsilon)\,\varepsilon d\varepsilon \quad \text{for } n \text{ odd,}$$

$$= \int_0^1 \varphi(\varepsilon)\, \varepsilon d\varepsilon \quad \text{for } n \text{ even.}$$

In applying these formulæ, attention must be paid to the several values of $\varphi'(\varepsilon)$ and the corresponding limits falling between those given in the above integrals. They give

$$\left.\begin{aligned}
\varepsilon_p &= \frac{2n+1}{2n} - \frac{1}{2n}\sqrt{[2n(n-1)]} \\
\varepsilon_m &= \sqrt{\left(\frac{(2n-1)^4-1}{96n^3(n-1)}\right)} \\
\varepsilon_a &= \frac{(2n-1)^3-1}{24n^2(n-1)}
\end{aligned}\right\} n \text{ odd}; \qquad (6)$$

$$\left.\begin{aligned}
\varepsilon_p &= 1 - \tfrac{1}{2}\sqrt{2} \quad \text{for } n = 2 \\
\varepsilon_p &= \frac{2n-1}{2n} - \frac{1}{2n}\sqrt{[2n(n-1)-1]} \\
&\quad \text{for } n > 2 \\
\varepsilon_m &= \sqrt{\left(\frac{n^4+(n-1)^4-1}{12n^3(n-1)}\right)} \\
\varepsilon_a &= \frac{n^3+(n-1)^2-1}{6n^2(n-1)}
\end{aligned}\right\} n \text{ even.} \qquad (7)$$

For obvious reasons these formulæ, for n odd and even, should give identical results for ε_p, ε_m, and ε_a, when $n = \infty$; moreover these results should be the same as those for $n = 2$. Such is the case. The value $n = \infty$ is not permissible, however, in these formulæ when applied to such tables as we use, since the tabular differences are always finite. To give consistent results when applied to such tables, n must be restricted to values not exceeding twice the greatest tabular difference. For values of n greater than this limit, the probable, mean, and average errors will be less than those given by (6) and (7) and will approach those for the system expressed by equation (1) as n approaches infinity or unity. The process outlined above may be readily applied to determine the laws of frequency of ε when n exeeeds the limit just mentioned.

As to the more general case, in which the interpolating factor t is a proper fraction of the form m/n, m and n being integers, a moment's consideration will show that it is comprised practically, or so far as numerical values are concerned, in the first case where $t = 1/n$. For, supposing m and n to have no common factor, the possible values of ε_3 corresponding to m/n are identical with those for $1/n$. Thus for $m/n = \frac{3}{7}$ the values of ε_3 are the same as those for $m/n = \frac{1}{7}$; those for $\frac{2}{10}$ and $\frac{4}{10}$ the same as those for $\frac{1}{5}$; those for $\frac{3}{10}$ the same as those for

$\frac{1}{10}$, and so on. The values of $\varphi(\varepsilon_0)$ and hence $\varphi(\varepsilon)$ for $t = m/n$ will indeed differ from those when $t = 1/n$. $\varphi(\varepsilon)$ will be of a somewhat more complex character.* But the limiting values of ε are the same in the two cases, and hence the probability functions cannot give probable, mean, or average errors differing materially in numerical value from those when $t = 1/n$. The restriction mentioned above relative to the magnitude of n for a given table must be borne in mind.

The following table gives the probable, mean, average, and maximum actual errors corresponding to the interpolating factor $t = 1/n$ for values of n from 1 to 10. The values of these errors for $n = 1$, or for a tabular value, are given in the first line for the sake of ready comparison.

Interpolating Factor.	Probable Error.	Mean Error.	Average Error.	Maximum Actual Error.
1	0.250	0.289	0.250	$\frac{1}{2}$
$\frac{1}{2}$	0.292	0.408	0.333	1
$\frac{1}{3}$	0.256	0.347	0.287	$\frac{5}{6}$
$\frac{1}{4}$	0.276	0.382	0.313	1
$\frac{1}{5}$	0.268	0.370	0.303	$\frac{9}{10}$
$\frac{1}{6}$	0.277	0.385	0.315	1
$\frac{1}{7}$	0.274	0.380	0.311	$\frac{13}{14}$
$\frac{1}{8}$	0.279	0.389	0.318	1
$\frac{1}{9}$	0.278	0.386	0.316	$\frac{17}{18}$
$\frac{1}{10}$.	0.281	0.392	0.320	1

The mean of the average errors in the above table for $t = \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}$, and $\frac{5}{10}$, using $\frac{1}{10}$ and $\frac{1}{5}$ as arguments in the place of $\frac{3}{10}$ and $\frac{4}{10}$, is 0.316. This number coincides closely, as it should, with the average of the 1000 actual errors given by Prof. Howe in No. 6 of the ANNALS.

---

*Thus, for example, if $t = \frac{2}{5}$ the probability function is

$$\begin{aligned}\phi(\varepsilon) &= 1 \text{ for values of } \varepsilon \text{ between } 0 \text{ and } \pm \tfrac{1}{10},\\ &= \tfrac{5}{6}\left(\tfrac{13}{10} \pm \varepsilon\right) \text{ for valûes of } \varepsilon \text{ between } \pm \tfrac{1}{10} \text{ and } \pm \tfrac{3}{10},\\ &= \tfrac{5}{3}\left(\tfrac{4}{5} \pm \varepsilon\right) \text{ for values of } \varepsilon \text{ between } \pm \tfrac{3}{10} \text{ and } \pm \tfrac{7}{10},\\ &= \tfrac{5}{6}\left(\tfrac{9}{10} \pm \varepsilon\right) \text{ for values of } \varepsilon \text{ between } \pm \tfrac{7}{10} \text{ and } \pm \tfrac{9}{10};\end{aligned}$$

and the probable, mean, and average errors are 0.261, 0.351, and 0.290 respectively.